DOCUMENT RESUME

ED 079 702 CS 000 655

AUTHOR Felsenthal, Helen
TITLE Readability: Computer Utilization.
PUB DATE May 73
NOTE 11p.; Paper presented at the Annual Meeting of the International Reading Assn. (18th, Denver, May 1-4, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Childrens Books; *Computer Programs; History Textbooks; *Readability; Reading Interests; *Reading Level; Reading Material Selection; *Reading Research; *Textbook Evaluation
IDENTIFIERS *Textual Analysis of Language Samples

ABSTRACT

Textual Analysis of Language Samples (TEXAN) is a computer program which can count a number of variables needed for use in readability formulas. Three studies which utilize TEXAN are reported in this paper: (1) In 1972, Norman and Helen Felsenthal randomly selected 20 books from the 1306 in Eakin's "Good Books for Children" and calculated their internal consistency using Gunning, Spache, and two Flesch formulas. The results disprove the speculation that difficulty increases from beginning to end in many children's books. The study also compared the four readability estimates; correlations varied from high to negative. (2) In a 1973 study, Alden J. Moe investigated the readability of selected Newbery Award Books. Fry, Gunning, and Lorge formulas were used but did not provide the same grade level estimates when applied to a single sample. (3) Also in 1973, Norman Felsenthal analyzed the readability and specialized vocabulary of nine selected U. S. history texts in grades 5, 8, and 11 using Flesch, Fry, and Lorge formulas. Results indicated the three fifth-grade books were in excess of their intended level of usage. The others were closer to their intended levels but a wide variation in scores among the three formulas existed. Other studies using computer programs are also reviewed. (TO)

ED 079702

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Helen Felsenthal
Education Department
Language Arts/Reading
Purdue University
West Lafayette, Indiana 47907

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Helen Felsenthal
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER"

# READABILITY: COMPUTER UTILIZATION

Annual Meeting, International Reading Association
Denver, Colorado
1973

Session: Wednesday, May 2, 10:45 a.m. - 11:45 a.m.

Matching child with book in terms of interest and readability level has been a long-time problem of the classroom teacher. Interest level and readability are closely related because the desire to read is invariably reduced when material is too difficult.

Although much research has centered on readability, the classroom teacher still has very little information concerning the readability of materials, especially trade books. Reference periodicals usually give a gross estimate such as "for use in intermediate grades, or for use in grades 3-5". Consequently, the teacher is often left with the task of determining a more precise difficulty level.

The main reason for this limited information is not the lack of formulae to identify readability but rather the tedious and time consuming work which is necessary to collect the data needed to calculate readability.

CS 000 655

This paper describes a computer program which greatly reduces the time involved in the calculation of readability and also reviews several studies which have utilized computer programs to determine readability. Although teachers seldom have access to computers, the technique described can be used by publishing companies and/or curricula specialists and the information made available to teachers.

The term 'readability' has been defined and interpreted many ways. Edgar Dale and Jeanne Chall, two of the best known researchers in this speciality, state: "In the broadest sense, readability is the sum total (including interactions) of all those elements within a given piece of printed material that affects the success which a group of readers have with it. The success is the extent to which they understand it, read it at optimum speed and find it interesting (1948, p. 38)."

The matching of interest and readability is emphasized in Gilliland's statement:

> On the one hand there is a collection of individuals with given interests and reading skills. On the other hand, there is a range of books and other reading materials, differing widely in content, style and complexity. The extent to which books can be read with profit will be determined largely by the way in which the two sides are matched. For example, a person who is a competent reader may soon be deterred from reading if her choice is restricted to simple, repetitive texts. Similarly, a person with limited reading ability may soon become discouraged if he is given texts which are beyond his comprehension (1972, p. 12).

A concise and inclusive definition is offered by Lamb (1973): "Readability is the sum of factors, and the interactive effect of these factors, which may be greater than the sum, affecting an individual's ability to comprehend what is read. Factors typically considered in readability are number of words in a sentence, number of syllables in words, and frequently, an analytical comparison of the words in a selection with those included on a standardized list of some type" (p.2).

Interest in readability dates as far back as the McGuffey readers where an attempt was made to grade materials in terms of difficulty level. Lamb (1973), in her recent review of literature concerning readability, notes that Thorndike's Teachers Word Book (1921) was one of the first efforts to objectively measure the difficulty level of reading materials. Virtually all of the early attempts to analyze readability relied upon vocabulary variables as the main factors in determining difficulty level. Between 1934 and 1938 readability research efforts were broadened and more than vocabulary factors were considered. Reliance on word lists compiled by Thorndike and others diminished and attention was given to factors such as sentence length and syntactical construction such as parts of speech. Efforts to make readability formulas more efficient by reducing the number of variables in the formulas was noted during the 1940's and 1950's (Dale and Chall, 1948; Flesch, 1948; Lorge, 1939; Yoakum, 1951).

Letter redundancy ( Carterette and Jones, 1963), and independent clause frequency (Strickland, 1962) are two additional variables included in the more recent studies. The Cloze procedure, a patterned deletion of words from passages, has also been utilized with considerable interest by researchers in readability (Bormuth, 1963, 1966; Klare, 1963; Ramanauskas, 1972; Rankin and Culhane, 1969; Taylor, 1953).

The attempt to automate readability measurement is of more recent origin and has, for the most part, utilized computer programming. One notable exception is the Readability Index Tabulator which is attached to an electric typewriter to collect readability data. The attachment, developed by Smith and Kincaid (1970) tabulates the number of strokes (letters), the number of words, and the number of sentences. Information from the tabulator is then utilized using a computer to determine

readability (Kincaid et. al, 1972).

Computer usage eliminates the tedious tasks of counting words, syllables, sentences, and other variables which are needed for use in readability formulas. One such program, TEXAN (Textual Analysis of Language Samples) has been developed at Purdue University. This program counts a number of variables including: total words (any combination of alphanumeric characters delineated by spaces), total non-exempt words (those words not included in a special listing such as the Dale list of 769 easy words), different non-exempt words, special words (those words, up to 100, which the programmer designates; e.g., a list of pronouns or other words with special significance), statements, questions, exclamations, total sentences, quotations, words per sentence, non-exempt words per sentence, characters per word, characters per non-exempt word, and average occurrences of each non-exempt word (listed either alphabetically, by frequency, or by first occurrence as determined by the sub-routine requested by the user).

One key element of many readability formulas which can not be obtained directly by the TEXAN program is syllable count. This can be accurately estimated, however, by dividing the number of letters in any message by 3.1127. This constant is derived from data analyzed in a research study where forty language samples were studied and correlations were run between a man-made syllable count and both letter count and vowel count. Number of characters and number of syllables correlated at .98 with a mean ratio of characters to syllables of 3.1127. Number of vowels and number of syllables correlated at .96 with a mean ratio of vowels to syllables of 1.761 (Felsenthal, Shamo, Bittner, 1971).

Three studies which have utilized the TEXAN are reported below. Felsenthal and Felsenthal (1972) calculated readability internal consistency of 20 books randomly selected from among the 1306 books included in Eakin's (1962) Good Books for Children. Each book yielded three language samples of approximately 200-300 words each from the first third of the book, the middle portion, and the final portion. The sixty language samples were key punched and processed utilizing the TEXAN program. Measurements were transferred to work sheets and a calculator was used to ascertain the readability for each passage using four different readability formulas: Gunning's Fog Index (Gunning, 1952), Spache (Spache, 1957), Flesch Reading Ease (Flesch, 1963) and Flesch's Human Interest (Flesch, 1963). Readability indices for the three samples from each book were compared and Chi-square procedures were employed to determine if readability variations within each book were excessive. None of the eighty Chi-squares (four indices x twenty books) was significant. Internal consistency of the twenty books as a single sample was measured in the second analysis. One-way analyses of variance were performed for each of four readability indices. None of the four indicated significant differences between the first, second, and third portions of the books. This study verified the internal consistency of these books which disproves the speculation by some educators that books become harder as they progress from beginning to end.

A secondary purpose of this study was to compare the readability estimates of the four readability indices utilized. The correlations indicated a relatively high correlation between the Gunning and Spache formulas, a moderately low correlation between the two Flesch formulas, and a negative correlation between the two Flesch formulas and the Gunning and Spache grade level indicators. This negative correlation

was anticipated since higher scores on the two Flesch indices mean greater ease in reading as opposed to grade level indices which increase in value as they increase in difficulty.

In another study, Moe (1973) investigated the readability of selected Newbery Award Books and developed a word list of 200 high frequency words common to all books analyzed. Five 100 word samples were taken from each of three parts of the books: the middle of the first chapter, the middle of the middle chapter, and the middle of the last chapter. Readability was estimated using three formulas: Fry (1968), Gunning (1952), and Lorge (1959). Results indicated that the three readability formulas usually did not provide the same grade level estimates when applied to a single sample; however, the Lorge and the Fry provided similar results. The Gunning generally rated samples as being more difficult than either the Lorge or Fry estimates. In general, the sample of Newbery Award Books which were analyzed in this study had readability levels primarily in the fifth through seventh grade levels although the range was from second grade through ninth grade.

The 200 high frequency words were identified by analyzing all 75 language samples (three passages from each of 25 books). The particular words can be found in the research reported by Moe (1973), however, the first eleven words were the same as those identified in an earlier word study of primary-grade trade books (Moe, 1972). These words were: the, and, a, to, he, of, in, was, his, it, and I.

Another research study (Felsenthal, 1973) analyzed the readability and specialized vocabulary of selected U.S. history texts in grades five, eight, and eleven, and made comparisons between materials designed for each of these three grade levels. Data from a fourth or "news

periodical" level (Time, U.S. News and World Report, and Newsweek) was also examined and described.

Nine U.S. history textbooks currently in heavy use throughout the country were selected to provide data for the study. Three of these texts were specifically written for and are widely used in fifth grade classes; three other texts are used in eighth grade and the remaining three in eleventh grade. Five 200 to 300 word language samples were randomly selected from each of the nine books for a total of forty-five language samples. An additional fifteen samples were drawn from news magazines; five each from the magazines previously mentioned. The total sample consisted of sixty separate selections and total data exceeded 14,000 words. The sixty language samples were key-punched and processed utilizing the TEXAN program.

The author identified readability through the use of three formulas: Flesch (1963), Fry (1968), and Lorge (1959). Results indicated that all three fifth grade textbooks yielded readability scores in excess of their intended levels of usage (i.e. from one and a half to two years higher than the designated fifth grade level). Readability scores for the eighth and eleventh grade books were closer to their intended level however, although there was wide variation among the three formulas.

Readability scores for the news periodicals seemed closely related to those of the eleventh grade texts. The greatest factor of change across the fifth, eighth, and eleventh grade levels was in the number and percentage of large words. News periodicals, however, used fewer large words than did eleventh grade texts. The one factor that showed a consistent rate of increase across all four levels was sentence

length with fifth, eighth, eleventh, and the news periodical selections running approximately 12, 14, 18, and 22 words per sentence respectively.

Another intent of the research, the development of a hierarchy of specialized words related to the study of U.S. history and current events, was not realized. A special count of those words used four or more times and not in the Dale list of 769 easy words yielded a paucity of social studies words. In the fifth grade sample only "slavery" could be identified as a word unique to social studies. "Armistice" was the only unique word found in the eighth grade samples. In the eleventh grade sample only "federal", "political", "representative", and "Republican" could be labeled "social studies" words. Even more surprising was the absence of frequently used social words in the news periodicals. Only six words were unique ("American", "history", "U.S.", "Johnson(s)", "President(ial)", and "Vietnam"), and the latter three were a function of the particular date and time of the sample.

As stated earlier all three of the studies previously cited utilized the TEXAN computer program to identify the variables needed for readability identification. In each case the TEXAN program generated much more data than was actually used by the researchers. Consequently, additional analyses may be performed at a later date.

In summary, conclusions from the three studies using TEXAN reveal:

1) Trade books tend to be internally consistent in terms of readability. Classroom teachers need not be concerned that the reading level becomes more difficult, or changes at all, as the student progresses through a book.

2) Selected social studies texts used for U.S. history in grades 5, 8, 11 tend to have a greater reading difficulty level than the assigned grade. This is particularly true of fifth grade texts where the readability level was $1\frac{1}{2}$-2 years higher than the grade. Therefore, the readers of these textbooks should be at least average or preferably above average readers. Corrective and remedial readers cannot be expected to use these books.

3) It is both difficult to ascertain and unnecessary to teach a particular social studies vocabulary since those words utilized in social studies texts are virtually identical to words employed in standard reading.

4) News periodicals such as Time, U.S. News and World Report and Newsweek can be used as supplementary reading at the eleventh grade level since the readability of these magazines approximates that of the eleventh grade social studies texts.

5) The readability level of such important books as the Newbery Award and other popular books should be more precisely identified for readability. In the past most Newbery Award books have readability levels between fifth and seventh grade.

6) A list of high frequency words (such as the Dolch) can be determined for various language samples. It is important that just eleven words (the, and, a, to, he, of, in, was, his, it, and I) appear to be used extensively in almost all language samples.

Some implications can also be drawn from the three computer-assisted studies:

1) Although there are many readability formulas currently in use, few hold their value across a broad range of difficulties. The Spache formula is best limited to primary usage, Lorge seems appropriate for junior high, and Fry for high school. The Flesch Reading Ease Index seems to have the broadest range of the formulas examined in the three studies.

2) More readability checks should be made on texts used in the content areas such as social studies and science. The actual difficulty of content subject texts may be quite a bit higher than the designated grade level.

3) Since content texts are often authored by more than one person, the consistency of readability throughout the text should be examined. This process does not seem to be necessary for single-authored trade books, however.

It is indeed feasible to utilize a computer program to measure stylistic variables and calculate readability levels. This automation of a previous tedious task promises to offer the classroom teacher much more information concerning the readability of various materials.

REFERENCES

Bormuth, J. R. "Cloze as a Measure of Readability," in International Reading Association Conference Proceedings, 8:131-4, 1963.

Bormuth, John. "Readability: A New Aprroach," The Reading Research Quarterly, 1:81-83, 1966.

Carterette, E.C., and Jones, Margaret. "Redundancy in Children's Texts," Science, 140:1309-1311, 1963.

Dale, Edgar and Jeanne Chall. "A Formula for Predicting Readability," Educational Research Bulletin, 27: 11-20, 37-54, 1948.

Eakin, Mary K. Good Books for Children. Chicago: University of Chicago Press, 1962.

Felsenthal, Norman. "Measuring the Readability and Specialized Vocabulary of Selected Social Studies Texts," Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, February, 1973.

Felsenthal, Norman and Helen Felsenthal. "Utilizing the Computer To Assess the Readability of Language Samples," Paper presented at the Annual Meeting of the American Education Research Association, Chicago, April, 1972.

Felsenthal, Norman, Shamo, G. Wayne, and John R. Bittner. "A Comparison of Award-Winning Radio Commercials with their Day-to-Day Counterparts," Journal of Broadcasting, XV (summer, 1971), 309-315.

Flesch, R. "A New Readability Yardstick," Journal of Applied Psychology, 32: 221-233, 1948.

Flesch, Rudolf. How To Write, Speak and Think More Effectively. New York: Signet, 1963, 298-302, and 303-306.

Fry, E. A. "A Readability Formula That Saves Time," Journal of Reading, 2: 513-16, 1968.

Gilliland, John. "Readability," London University of London Press, Ltd., 1972, p. 12.

Gunning, Robert. The Techniques of Clear Writing. New York: McGraw-Hill, 1952.

Kincaid, J. Peter, Van Deusen, John, Thomas, Georgelle, Lewis, Robert, Anderson, Patricia Tanner and Linda Moody. Use of the Automated Readability Index for Evaluating Peer-Prepared Material for Use In Adult Reading Education. U.S. Department of Health, Education, and Welfare, Office of Education, September, 1972.

Klare, George R. The Measurement of Readability. Ames, Iowa: The Iowa State University Press, 1963, pp. 75-80.

Lamb, Pose. "The Measurement of Readability: Review of Literature," Paper presented at the Annual Meeting of the American Education Research Association, New Orleans, Louisiana, 1973.

Lorge, I. "Predicting Reading Difficulty of Selections for Children," Elementary English Review, 16: 141-142, 1939.

Moe, Alden J. High-Frequency Words: Word Cards for Beginning Readers. St. Paul: Ambassador Publishing, 1972.

Moe, Alden J. "The Readability of Newbery Award Books: A Computer-Assisted Analysis," Paper presented at the Annual Meeting of the American Education Research Association, New Orleans, Louisiana, February, 1973.

Moe, Alden J. "Vocabularies for Beginning Readers: A Computer-Assisted Analysis of Trade Books for Young Children," Paper presented at the Annual Meeting of the International Reading Association, Denver, Colorado, 1973.

Ramanauskas, Sigita. "The Responsiveness of Cloze Readability Measures to Linguistic Variables Operating Over Segments of Text Larger Than a Sentence," Reading Research Quarterly, 8: 72 88, 1972.

Rankin, E. F. and Culhane, J.W. "Comparable Cloze and Multiple Choice Comprehension Test Scores," Journal of Reading, 13 193-198, 1969.

Smith, Edgar A. and J. Peter Kincaid. "Derivation and Validation of the Automated Readability Index for Use with Technical Materials," Human Factors, 12 (5), 1970, 457-464.

Spache, George. "A New Readability Formula for Primary-Grade Reading Materials," Elementary School Journal, XLI (Fall, 1957), 214-217.

Strickland, Ruth. The Language of Elementary School Children. Bulletin of the School of Education, 1962.

Taylor, W. L. "Cloze Procedure: A New Tool for Measuring Readability," Journalism Quarterly, 30: 415-33, 1953

Thorndike, Edward L. The Teacher's Word Book. New York: Teacher's College, Columbia University, 1921.

Yngve, V. H. "Computer Programs for Translation," Scientific American, 206: 68-76, 1962.

Yoakum, G. A. "Determining the Readability of Instructional Materials" Report of the Seventh Annual Conference on Reading, Pittsburgh: University of Pittsburgh, 47-53, 1951.